

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/533,361  
Source: PT/10  
Date Processed by STIC: 5/12/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:34

Input Set : A:\44-05sequence listing.txt

Output Set: N:\CRF4\05122005\J533361.raw

3 <110> APPLICANT: BURTON, KERRY  
 4 CHALLEN, MICHAEL  
 5 ELLIOTT, TIMOTHY  
 6 SREENIVASAPRASAD, SURAPAREDDY  
 7 EASTWOOD, DANIEL  
 8 MOLLOY, SHANNON  
 10 <120> TITLE OF INVENTION: SELECTIVE EXPRESSION IN FILAMENTOUS FUNGI  
 12 <130> FILE REFERENCE: 44-05  
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/533,361  
 C--> 14 <141> CURRENT FILING DATE: 2005-04-29  
 14 <150> PRIOR APPLICATION NUMBER: PCT/GB2003/004716  
 15 <151> PRIOR FILING DATE: 2003-10-31  
 17 <150> PRIOR APPLICATION NUMBER: UK 0225390.4  
 18 <151> PRIOR FILING DATE: 2002-10-31  
 20 <160> NUMBER OF SEQ ID NOS: 37  
 22 <170> SOFTWARE: PatentIn version 3.1  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 1736  
 26 <212> TYPE: DNA  
 27 <213> ORGANISM: Agaricus bisporus  
 29 <220> FEATURE:  
 30 <221> NAME/KEY: CDS  
 31 <222> LOCATION: (19)..(1674)  
 32 <223> OTHER INFORMATION:  
 W--> 35 <400> 1  
 36 ggctgagctc tattcatc atg gcg tgc gaa cga cag att gaa gaa ctt ccc 51  
 37 Met Ala Ser Glu Arg Gln Ile Glu Glu Leu Pro  
 38 1 5 10  
 40 agt tat aag tac gcc tat att ttg acg gca tgc gct tgc ttg gga agt 99  
 41 Ser Tyr Lys Tyr Ala Tyr Ile Leu Thr Ala Ser Ala Cys Leu Gly Ser  
 42 15 20 25  
 44 gtg ttt cac ggg tgg gat gta ggc ctt ata gga ggc ata ctc tcg cta 147  
 45 Val Phe His Gly Trp Asp Val Gly Leu Ile Gly Gly Ile Leu Ser Leu  
 46 30 35 40  
 48 cgg tca ttt caa gaa tat ctc ggg atc aat aca aaa aat gcc gtc aag 195  
 49 Arg Ser Phe Gln Glu Tyr Leu Gly Ile Asn Thr Lys Asn Ala Val Lys  
 50 45 50 55  
 52 aaa gcg att cta gac gga aac atc atc tct gtg ctc caa gcc gga tgt 243  
 53 Lys Ala Ile Leu Asp Gly Asn Ile Ile Ser Val Leu Gln Ala Gly Cys  
 54 60 65 70 75  
 56 ttt ttt ggc gcg ctt gga acc gga tat ctc tct agt cga ttc ggc cga 291  
 57 Phe Phe Gly Ala Leu Gly Thr Gly Tyr Leu Ser Ser Arg Phe Gly Arg  
 58 80 85 90

P.6

## RAW SEQUENCE LISTING

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:34

Input Set : A:\44-05sequence listing.txt

Output Set : N:\CRF4\05122005\J533361.raw

```

60 aga ccc tgt ctt att gca tct ggt att gtg tat ata act ggc ggt ttg      339
61 Arg Pro Cys Leu Ile Ala Ser Gly Ile Val Tyr Ile Thr Gly Gly Leu
62          95          100          105
64 ctg caa tgc act gtc ggt ttg gga ccc tcg caa gct gct gct cta cac      387
65 Leu Gln Cys Thr Val Gly Leu Gly Pro Ser Gln Ala Ala Ala Leu His
66          110          115          120
68 gtg ttc tat att ggc agg ttc att tct ggt atc ggt gtt ggg atg gtg      435
69 Val Phe Tyr Ile Gly Arg Phe Ile Ser Gly Ile Gly Val Gly Met Val
70          125          130          135
72 tcc act ctc gtg cct ttg tat att tcg gag tgt gtc cct agg act ata      483
73 Ser Thr Leu Val Pro Leu Tyr Ile Ser Glu Cys Val Pro Arg Thr Ile
74 140          145          150          155
76 cgc ggg cgc tgt act gga aca ctc caa ttt gcg act aac agt ggt ctg      531
77 Arg Gly Arg Cys Thr Gly Thr Leu Gln Phe Ala Thr Asn Ser Gly Leu
78          160          165          170
80 atg ctg ggc ttc tgg gtc aac tac agc gtg tcg aaa aac gtg ccc ttt      579
81 Met Leu Gly Phe Trp Val Asn Tyr Ser Val Ser Lys Asn Val Pro Phe
82          175          180          185
84 ggt gaa atg caa tgg cga att ccg tta att atc caa atg att ccg agc      627
85 Gly Glu Met Gln Trp Arg Ile Pro Leu Ile Ile Gln Met Ile Pro Ser
86          190          195          200
88 ctc ttg ttc atc ata gcc atg ttt ttc caa cca gaa tcg ccg aga tgg      675
89 Leu Leu Phe Ile Ile Ala Met Phe Phe Gln Pro Glu Ser Pro Arg Trp
90          205          210          215
92 ctt gtt gaa cac ggg aaa cac aag gaa gct gcg acg gta ctg gcg cgt      723
93 Leu Val Glu His Gly Lys His Lys Glu Ala Ala Thr Val Leu Ala Arg
94 220          225          230          235
96 act ggc ggc aag gat gtt gat cat cct agt gtt gta cag aca ctg gag      771
97 Thr Gly Gly Lys Asp Val Asp His Pro Ser Val Val Gln Thr Leu Glu
98          240          245          250
100 gag atc aag caa gaa ttt gtg gcg agt aaa caa cca tcg ttt tta aag      819
101 Glu Ile Lys Gln Glu Phe Val Ala Ser Lys Gln Pro Ser Phe Leu Lys
102          255          260          265
104 cag att cgc ctg gtc ggt gaa tcg agg gct gtt gcc ctg agg tgc ttt      867
105 Gln Ile Arg Leu Val Gly Glu Ser Arg Ala Val Ala Leu Arg Cys Phe
106          270          275          280
108 ata cca ccg cta gtg atg ttc ttc cag cag tgg acg ggt aca aat gcc      915
109 Ile Pro Pro Leu Val Met Phe Phe Gln Gln Trp Thr Gly Thr Asn Ala
110          285          290          295
112 atc aac ctt tat agt ccc gaa gta ttc cgt cat ctt gga atc cat ggc      963
113 Ile Asn Leu Tyr Ser Pro Glu Val Phe Arg His Leu Gly Ile His Gly
114 300          305          310          315
116 acc agc ggg gct ctc ttc gct act ggt gtt tat ggc gtg gtg aag gtt      1011
117 Thr Ser Gly Ala Leu Phe Ala Thr Gly Val Tyr Gly Val Val Lys Val
118          320          325          330
120 gtt tca gtt gca ctt gcc ctc act ttt gct gtc gaa cgc ttt gga cgc      1059
121 Val Ser Val Ala Leu Ala Leu Thr Phe Ala Val Glu Arg Phe Gly Arg
122          335          340          345
124 aag aga ggg ttg att ttt ggt ggt atc ggc caa gca ctt atg atg ttt      1107

```

## RAW SEQUENCE LISTING

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:34

Input Set : A:\44-05sequence listing.txt

Output Set: N:\CRF4\05122005\J533361.raw

```

125 Lys Arg Gly Leu Ile Phe Gly Gly Ile Gly Gln Ala Leu Met Met Phe
126          350          355          360
128 tgg ttg gga ggt tat agt gcc acc cac caa gac ggt act gtc agt cct      1155
129 Trp Leu Gly Gly Tyr Ser Ala Thr His Gln Asp Gly Thr Val Ser Pro
130          365          370          375
132 gcg agt cat gtt tcc att gtt gca ctc tac ttg tat ggt gca ttc ttc      1203
133 Ala Ser His Val Ser Ile Val Ala Leu Tyr Leu Tyr Gly Ala Phe Phe
134 380          385          390          395
136 tcc atg gga tgg ggc cca tta cca tgg gtc gtc gct gga gag gtt gca      1251
137 Ser Met Gly Trp Gly Pro Leu Pro Trp Val Val Ala Gly Glu Val Ala
138          400          405          410
140 cct aac cat gtc cgc tcc ttc gcc ctc tcc atc gcc gtt gga act cat      1299
141 Pro Asn His Val Arg Ser Phe Ala Leu Ser Ile Ala Val Gly Thr His
142          415          420          425
144 tgg ctc ttc ggg ttt gtg ata tca aaa gtg acg cca att atg ttg gac      1347
145 Trp Leu Phe Gly Phe Val Ile Ser Lys Val Thr Pro Ile Met Leu Asp
146          430          435          440
148 cgt atc aaa tat ggc aca ttc cta ctc ttc gga ttc tgt tgc atg ata      1395
149 Arg Ile Lys Tyr Gly Thr Phe Leu Leu Phe Gly Phe Cys Cys Met Ile
150          445          450          455
152 gta gcg aca tgg gct tat ttc tgt cta cct gag aca agt ggg ttc gct      1443
153 Val Ala Thr Trp Ala Tyr Phe Cys Leu Pro Glu Thr Ser Gly Phe Ala
154 460          465          470          475
156 ctg gag gac atc aaa tat ctg ttc gag cga gac gtc atc att cgt tca      1491
157 Leu Glu Asp Ile Lys Tyr Leu Phe Glu Arg Asp Val Ile Ile Arg Ser
158          480          485          490
160 ttg cag gac gct ccc ggt gga aaa ata ttc ttg ggg ggg agg cgt gtg      1539
161 Leu Gln Asp Ala Pro Gly Gly Lys Ile Phe Leu Gly Gly Arg Arg Val
162          495          500          505
164 gaa tct gta gct tcg ttg aaa gag agg cgc gtt gga gtc gct ggt gag      1587
165 Glu Ser Val Ala Ser Leu Lys Glu Arg Arg Val Gly Val Ala Gly Glu
166          510          515          520
168 cag ggt gag aag ata act ggt cta aat tcg gaa ttg gaa gat gtt tcc      1635
169 Gln Gly Glu Lys Ile Thr Gly Leu Asn Ser Glu Leu Glu Asp Val Ser
170          525          530          535
172 tca aaa aaa tca aca ttg aag gaa act tca tcc gtt tga tatatagtct      1684
173 Ser Lys Lys Ser Thr Leu Lys Glu Thr Ser Ser Val
174 540          545          550
176 ccaaattcta ttgtaatgcc attttcccaa ttcaaaaaaaaa aaaaaaaaaa aa      1736
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 551
181 <212> TYPE: PRT
182 <213> ORGANISM: Agaricus bisporus
184 <400> SEQUENCE: 2
186 Met Ala Ser Glu Arg Gln Ile Glu Glu Leu Pro Ser Tyr Lys Tyr Ala
187 1          5          10          15
190 Tyr Ile Leu Thr Ala Ser Ala Cys Leu Gly Ser Val Phe His Gly Trp
191          20          25          30
194 Asp Val Gly Leu Ile Gly Gly Ile Leu Ser Leu Arg Ser Phe Gln Glu

```

## RAW SEQUENCE LISTING

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:34

Input Set : A:\44-05sequence listing.txt

Output Set: N:\CRF4\05122005\J533361.raw

```

195          35          40          45
198 Tyr Leu Gly Ile Asn Thr Lys Asn Ala Val Lys Lys Ala Ile Leu Asp
199          50          55          60
202 Gly Asn Ile Ile Ser Val Leu Gln Ala Gly Cys Phe Phe Gly Ala Leu
203 65          70          75          80
206 Gly Thr Gly Tyr Leu Ser Ser Arg Phe Gly Arg Arg Pro Cys Leu Ile
207          85          90          95
210 Ala Ser Gly Ile Val Tyr Ile Thr Gly Gly Leu Leu Gln Cys Thr Val
211          100          105          110
214 Gly Leu Gly Pro Ser Gln Ala Ala Ala Leu His Val Phe Tyr Ile Gly
215          115          120          125
218 Arg Phe Ile Ser Gly Ile Gly Val Gly Met Val Ser Thr Leu Val Pro
219          130          135          140
222 Leu Tyr Ile Ser Glu Cys Val Pro Arg Thr Ile Arg Gly Arg Cys Thr
223 145          150          155          160
226 Gly Thr Leu Gln Phe Ala Thr Asn Ser Gly Leu Met Leu Gly Phe Trp
227          165          170          175
230 Val Asn Tyr Ser Val Ser Lys Asn Val Pro Phe Gly Glu Met Gln Trp
231          180          185          190
234 Arg Ile Pro Leu Ile Ile Gln Met Ile Pro Ser Leu Leu Phe Ile Ile
235          195          200          205
238 Ala Met Phe Phe Gln Pro Glu Ser Pro Arg Trp Leu Val Glu His Gly
239          210          215          220
242 Lys His Lys Glu Ala Ala Thr Val Leu Ala Arg Thr Gly Gly Lys Asp
243 225          230          235          240
246 Val Asp His Pro Ser Val Val Gln Thr Leu Glu Glu Ile Lys Gln Glu
247          245          250          255
250 Phe Val Ala Ser Lys Gln Pro Ser Phe Leu Lys Gln Ile Arg Leu Val
251          260          265          270
254 Gly Glu Ser Arg Ala Val Ala Leu Arg Cys Phe Ile Pro Pro Leu Val
255          275          280          285
258 Met Phe Phe Gln Gln Trp Thr Gly Thr Asn Ala Ile Asn Leu Tyr Ser
259          290          295          300
262 Pro Glu Val Phe Arg His Leu Gly Ile His Gly Thr Ser Gly Ala Leu
263 305          310          315          320
266 Phe Ala Thr Gly Val Tyr Gly Val Val Lys Val Val Ser Val Ala Leu
267          325          330          335
270 Ala Leu Thr Phe Ala Val Glu Arg Phe Gly Arg Lys Arg Gly Leu Ile
271          340          345          350
274 Phe Gly Gly Ile Gly Gln Ala Leu Met Met Phe Trp Leu Gly Gly Tyr
275          355          360          365
278 Ser Ala Thr His Gln Asp Gly Thr Val Ser Pro Ala Ser His Val Ser
279          370          375          380
282 Ile Val Ala Leu Tyr Leu Tyr Gly Ala Phe Phe Ser Met Gly Trp Gly
283 385          390          395          400
286 Pro Leu Pro Trp Val Val Ala Gly Glu Val Ala Pro Asn His Val Arg
287          405          410          415
290 Ser Phe Ala Leu Ser Ile Ala Val Gly Thr His Trp Leu Phe Gly Phe
291          420          425          430

```

## RAW SEQUENCE LISTING

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:34

Input Set : A:\44-05sequence listing.txt

Output Set: N:\CRF4\05122005\J533361.raw

```

294 Val Ile Ser Lys Val Thr Pro Ile Met Leu Asp Arg Ile Lys Tyr Gly
295         435                     440                     445
298 Thr Phe Leu Leu Phe Gly Phe Cys Cys Met Ile Val Ala Thr Trp Ala
299         450                     455                     460
302 Tyr Phe Cys Leu Pro Glu Thr Ser Gly Phe Ala Leu Glu Asp Ile Lys
303 465                     470                     475                     480
306 Tyr Leu Phe Glu Arg Asp Val Ile Ile Arg Ser Leu Gln Asp Ala Pro
307                     485                     -                     490                     495
310 Gly Gly Lys Ile Phe Leu Gly Gly Arg Arg Val Glu Ser Val Ala Ser
311                     500                     505                     510
314 Leu Lys Glu Arg Arg Val Gly Val Ala Gly Glu Gln Gly Glu Lys Ile
315                     515                     520                     525
318 Thr Gly Leu Asn Ser Glu Leu Glu Asp Val Ser Ser Lys Lys Ser Thr
319         530                     535                     540
322 Leu Lys Glu Thr Ser Ser Val
323 545                     550

```

326 &lt;210&gt; SEQ ID NO: 3

327 &lt;211&gt; LENGTH: 3638

328 &lt;212&gt; TYPE: DNA

329 &lt;213&gt; ORGANISM: Agaricus bisporus

331 &lt;220&gt; FEATURE:

332 &lt;221&gt; NAME/KEY: misc\_feature

333 &lt;222&gt; LOCATION: (3)..(3)

334 &lt;223&gt; OTHER INFORMATION: n is an unknown nucleotide

337 &lt;220&gt; FEATURE:

338 &lt;221&gt; NAME/KEY: misc\_feature

339 &lt;222&gt; LOCATION: (14)..(14)

340 &lt;223&gt; OTHER INFORMATION: n is an unknown nucleotide

343 &lt;220&gt; FEATURE:

344 &lt;221&gt; NAME/KEY: misc\_feature

345 &lt;222&gt; LOCATION: (510)..(510)

346 &lt;223&gt; OTHER INFORMATION: n is an unknown nucleotide

349 &lt;400&gt; SEQUENCE: 3

```

W--> 350 gtncgatggg ttentctggg ttaagttgca cgacgctttc ctttttcttt tatggcctgt      60
352 ctgccctttt aacgctttat ctttcggcag ccatggatgt ccttcgtcac cgtattatca      120
354 ctcttaatcg tgggtgtgag cacatggaaa ggttcattat cgtcatccca tgacgcggtg      180
356 caaaattcgt cattcagagt ggaaccgata ctaggagagg attttgaaaa ggctatcgta      240
358 tcttccgata cattcactcc aacgacatct atatcgggtt ctcaagctac gataaccggt      300
360 cctcacgcat cacattcttc tctgattgta atgaccgaag accgcggttc caccctaat      360
362 attcgaagag cttgttgcca atgtcaaaaa aggtcttttg gcaaaccggg atggccttat      420
364 gagcctccag ccccagcagc taacatcggg cagagaagga aaattcatcg ggcggttgaa      480
366 ttatcaccgt ttggttcctg agtcatctgn ragatgtacg cagatgggtga taccgtgttt      540
368 gattggcgcc gttggagaag aactatatta ttcgatggat tttttgttcg agtttgacac      600
370 agagacagag atgatatagg tttgctattg atgtagcaaa ggatcatttg acgatggcgc      660
372 atagggcgat ggttatcttt atgtctggaa ttataatatg tattgttccc cacttttctt      720
374 ttatatattat taataactaat tggaaagttc agttgttgga tgagcaaaagt tggtgagat      780
376 agaaactaga attcggattc ccataatctga ggtacctttt ccttccgctg gcaatcctgg      840
378 ccacttcgac gtggtgacgc agagggcgcg tgctattgtt agcacatgcc atawggatcg      900
380 acgttgctc tcgtacttcg cgctaggct cgctcatgcc tcgatgcac ttttcaattc      960

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/533,361

DATE: 05/12/2005  
TIME: 08:02:35

Input Set : A:\44-05sequence listing.txt  
Output Set: N:\CRF4\05122005\J533361.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 3,14,510  
Seq#:4; N Pos. 495  
Seq#:8; N Pos. 2774  
Seq#:10; N Pos. 1262  
Seq#:37; N Pos. 290,498,589

## VERIFICATION SUMMARY

DATE: 05/12/2005

PATENT APPLICATION: US/10/533,361

TIME: 08:02:35

Input Set : A:\44-05sequence listing.txt

Output Set : N:\CRF4\05122005\J533361.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No  
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:35 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:32  
L:350 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0  
M:341 Repeated in SeqNo=3  
L:501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:480  
L:564 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:6,Line#:561  
L:755 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:2760  
L:853 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:1260  
L:984 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:14,Line#:975  
L:990 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:14,Line#:981  
L:990 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:14,Line#:987  
L:1021 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:15,Line#:1018  
L:1328 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:240  
M:341 Repeated in SeqNo=37